

On the Grassmann condition number

Javier Peña* Vera Roshchina†

April 27, 2016

Abstract

We give new insight into the *Grassmann condition* of the conic feasibility problem

$$\text{find } x \in L \cap K \setminus \{0\}. \quad (1)$$

Here $K \subseteq V$ is a regular convex cone and $L \subseteq V$ is a linear subspace of the finite dimensional Euclidean vector space V . The Grassmann condition of (1) is the reciprocal of the *distance* from L to the set of *ill-posed* instances in the Grassmann manifold where L lives.

We consider a very general distance in the Grassmann manifold defined by two possibly different norms in V . We establish the equivalence between the Grassmann distance to ill-posedness of (1) and a natural measure of the *least violated* trial solution to the alternative to (1). We also show a tight relationship between the Grassmann and Renegar's condition measures, and between the Grassmann measure and a *symmetry* measure of (1).

Our approach can be readily specialized to a canonical norm in V induced by K , a prime example being the one-norm for the non-negative orthant. For this special case we show that the Grassmann distance ill-posedness of (1) is equivalent to a measure of the *most interior* solution to (1).

*Tepper School of Business, Carnegie Mellon University, USA, jfp@andrew.cmu.edu

†School of Science, RMIT University, AUSTRALIA, vera.roshchina@rmit.edu.au

1 Introduction

Assume V is a finite dimensional Euclidean vector space and $K \subseteq V$ is a closed convex cone. Given a linear subspace $L \subseteq V$, consider the feasibility problem

$$\text{find } x \in L \cap K \setminus \{0\}. \quad (2)$$

Feasibility problems of this form are pervasive in optimization. The constraints of linear, semidefinite, and more general conic programming are written in the form (2). The fundamental signal recovery property in compressive sensing can be stated precisely as the infeasibility of (2) for suitable K and L as it is nicely explained in [3].

We develop novel insights into the *Grassmann condition* of (2), which is defined as follows. Let $\text{Gr}(V, m)$ denote the Grassmann manifold of linear subspaces $L \subseteq V$ of dimension $m < \dim(V)$. Given $L \in \text{Gr}(V, m)$ consider the following measure of well-posedness of the problem (2): How far, in some suitable distance in $\text{Gr}(V, m)$, can L be perturbed without changing the feasibility status of (2). The size of the largest such perturbation of L is the *Grassmann distance to ill-posedness*. The reciprocal of this distance to ill-posedness is the *Grassmann condition* of L .

The Grassmann condition is a relatively recent development in the literature on conditioning for optimization initiated by Renegar [25, 26] and further studied by a number of authors [7, 10, 11, 14–16, 18–20, 22, 23]. In contrast to Renegar’s condition, which inherently depends on the *data* of some linear operator defining L , the Grassmann condition is data independent as it is entirely constructed in terms of objects in the Grassmann manifold. The initial developments on the Grassmann condition can be traced to the articles [1, 7, 13]. This Grassmann condition machinery has in turn been instrumental in further developments concerning the average behavior of condition numbers [2]. Our work is inspired by the main ideas in [1] and [13]. In spite of their conceptual overlap, the two articles [1, 13] were completed independently of each other. On the one hand, Amelunxen and Burgisser [1], extending ideas from Belloni and Freund [7], propose a Grassmann condition by working with a canonical projection-based distance in $\text{Gr}(V, m)$. They establish an interesting connection between the Grassmann and Renegar’s condition. They also give some characterizations of the Grassmann condition in terms of angles. On the other hand, Cheung, Cucker and Peña [13] introduce essentially the same concept of Grassmann condition albeit in a different format. They use a more general type of distance in the Grassmann manifold and a more nuanced type of ill-posedness but their

development is restricted to the special case $V = \mathbb{R}^n$, $K = \mathbb{R}_+^n$. Cheung et al. [13] establish several equivalences and bounds between the Grassmann condition and a variety of other condition measures for polyhedral feasibility problems.

This article combines and extends the main ideas and results from both [1] and [13]. Our approach applies to general regular cones as in [1] but we use a very general distance function in $\text{Gr}(V, m)$ in the same spirit of [13]. Section 2 below details the construction of the relevant distance function. Our first main result (Theorem 1) shows that when (2) is feasible, its Grassmann distance to ill-posedness matches a natural measure of the *least violated trial solution* to the alternative system to (2) namely

$$u \in L^\perp \cap K^* \setminus \{0\}. \quad (3)$$

Here L^\perp and K^* are the orthogonal complement of L and the dual cone of K respectively. We also give a counterpart of this result for the case when (2) is infeasible, that is, when (3) is feasible. This first result extends [13, Theorem 1] as well as [1, Proposition 1.6].

Our second main result (Theorem 2) establishes a tight connection between the Grassmann condition and Renegar's condition. In the particular case when the space L is defined via a linear isometry, Theorem 2 shows that the Grassmann and Renegar's condition measures are exactly the same. Otherwise Renegar's condition is sandwiched between the Grassmann condition and the Grassmann condition times the condition number of the linear operator defining L . The statement of Theorem 2 is nearly identical to [1, Theorem 1.4]. Our approach reveals that the relationship between Grassmann and Renegar's condition numbers indeed holds in much wider generality.

Our general approach can be seamlessly specialized to a canonical norm *induced* by K . Prime examples of this induced norm are the one-norm for the non-negative orthant and the nuclear norm for the cone of positive semidefinite matrices. When one of the norms in V is the norm induced by K , the Grassmann distance to ill-posedness of (2) can be characterized in terms of the *most interior* solution to (2). We also introduce a related family of *eigenvalue mappings* induced by K that allows us to relate the Grassmann condition with another popular condition measure, namely the *symmetry measure* which has been used by several authors [5, 6, 8, 15]. The exact connection between the symmetry measure and other condition measures has only been explored in the special case $V = \mathbb{R}^n$, $K = \mathbb{R}_+^n$ by Epelman and Freund [15]. We show that some of the results concerning the symmetry

measure established in [15] hold in more generality and are nicely related to the Grassmann condition measure (Theorem 3).

It should be noted that we pay a small price for the of generality of our approach. Our formal results are stated in terms of two related but different distances in the Grassmann manifold, as we detail in Section 2 below. This is necessary due to a certain asymmetry in these distance functions. The statements in Theorem 1 and Theorem 2 reflect this asymmetry: the claims concerning the distance to ill-posedness of (2) are slightly different in the feasible and infeasible cases. In the special case of Euclidean norms, this kind of asymmetry disappears and the relevant quantities from Theorem 1 can be equivalently phrased in terms of principal angles (Proposition 3). The particular Euclidean norm case also subsumes some of the main results from [1] within the context of our work.

We mention in passing the work by Lewis [22] and Lewis and co-authors [9, 14] that was also inspired by Renegar’s seminal work on condition numbers [25, 26]. The article [22] shows that Renegar’s distance to ill-posedness can be seen as a special case of the *distance to non-surjectivity* of sublinear set-valued mappings. The textbook [9] also includes a related discussion [9, Section 5.4]. The article [24] subsequently showed that the generalization in [22] is a special case of a generic and natural correspondence between conic systems and sublinear set-valued mappings. The article [14] extends the ideas in [22] to the more general concept of *radius of metric regularity* of set-valued mappings. Like Renegar’s approach, the underlying setting in [9, 14, 22] is inherently data-dependent in contrast to the data-independent approach of this article. Although a connection between our work and [9, 14, 22] might exist, the sharp difference in data dependence between the two approaches makes that potential connection far from evident.

The main sections of the paper are organized as follows. Section 2 presents our main technical developments. We construct two distances dist_G and $\overline{\text{dist}}_G$ in $\text{Gr}(V, m)$ determined by two arbitrary norms in V . We show (Theorem 1) the equivalence between the distance to ill-posedness of (2) and a measure of the trial solution that least violates its alternative (3). We also establish a tight relationship between the Grassmann condition and Renegar’s condition (Theorem 2). Section 3 specializes our development to some particular choices of norms in V , namely the Euclidean norm and a norm induced by K . For Euclidean norms, we show that our main construction and results can be stated in terms of *principal angles*. For the induced norm we show that the distance to ill-posedness coincides with a natural measure of the *most centered* solution to (2). Section 4 shows an additional connection between the Grassmann distance and the *symmetry* condition studied

by Belloni, Freund, and others [5, 6, 8, 15]. Finally Section 5 specializes our developments to two particularly important cases. The first one is the case when V is an Euclidean Jordan algebra and K is the cone of squares in V . The rich algebraic structure of this case in turn yields an interesting structure on the quantities characterizing the Grassmann condition and the symmetry condition. The second special case is $V = \mathbb{R}^n, K = \mathbb{R}_+^n$. The peculiar structure of this case enables us to refine our main developments by relying on the classical Goldman-Tucker partition theorem.

2 The Grassmann condition measure

Throughout the entire paper we assume V is a finite dimensional Euclidean vector space with inner product $\langle \cdot, \cdot \rangle$ and $K \subseteq V$ is a regular closed convex cone. We also assume that $\|\cdot\|, \|\!\|\!\cdot\!\|\|$ are norms in V , not necessarily the same, and neither of them necessarily Euclidean. We let $\|\cdot\|^*, \|\!\|\!\cdot\!\|\|^*$ denote the dual norms of $\|\cdot\|$ and $\|\!\|\!\cdot\!\|\|$ respectively. That is, for $u \in V$

$$\|u\|^* := \max_{\|x\|=1} \langle u, x \rangle, \quad \text{and} \quad \|\!\|u\!\|\|^* := \max_{\|\!\|x\!\|\|=1} \langle u, x \rangle.$$

Notice that by construction, the following *Hölder's inequality* holds for all $u, x \in V$

$$|\langle u, x \rangle| \leq \|u\|^* \cdot \|x\|, \quad |\langle u, x \rangle| \leq \|\!\|u\!\|\|^* \cdot \|\!\|x\!\|\|. \quad (4)$$

We let K^* denote the *dual cone* of K , that is,

$$K^* := \{u \in V : \langle u, x \rangle \geq 0 \forall x \in K\}.$$

Given a linear subspace $L \subseteq V$, let L^\perp denote the *orthogonal complement* of L in V , that is,

$$L^\perp := \{u \in V : \langle u, x \rangle = 0 \forall x \in L\}.$$

Our central objects of attention are the two feasibility problems (2) and (3) defined by a linear subspace $L \subseteq V$. Observe that (2) and (3) are *alternative systems*: one of them has a strictly feasible solution if and only if the other one is infeasible.

For ease of notation, we will use the following convention throughout the paper. The symbols v, x, z will denote “primal vectors”, that is, elements of L and K . On the other hand, the symbols u, w, y will denote “dual vectors”, that is, elements of L^\perp and K^* .

Define the distance functions $\text{dist}_G, \overline{\text{dist}}_G : \text{Gr}(V, m) \times \text{Gr}(V, m) \rightarrow \mathbb{R}$ as follows

$$\text{dist}_G(L_1, L_2) := \max_{\substack{x \in L_1 \\ x \neq 0}} \min_{v \in L_2} \frac{\|x - v\|}{\|x\|}, \quad \overline{\text{dist}}_G(L_1, L_2) := \max_{\substack{x \in L_1 \\ x \neq 0}} \inf_{\substack{v \in L_2 \\ v \neq 0}} \frac{\|x - v\|}{\|v\|}.$$

The infimum in the definition of $\overline{\text{dist}}_G(\cdot, \cdot)$ is essential. For instance, if $\|\cdot\| = \|\cdot\|_2$ and $L_1 \subseteq L_2^\perp$ then $\inf_{v \in L_2, v \neq 0} \frac{\|x - v\|_2}{\|v\|_2} = 1$ for all $x \in L_1, x \neq 0$ but this infimum is not attained.

The distances $\text{dist}_G(\cdot, \cdot)$ and $\overline{\text{dist}}_G(\cdot, \cdot)$ are not necessarily symmetric: In general, it may be the case that $\text{dist}_G(L_1, L_2) \neq \text{dist}_G(L_2, L_1)$ and $\overline{\text{dist}}_G(L_1, L_2) \neq \overline{\text{dist}}_G(L_2, L_1)$. For example, let $L_1, L_2 \in \text{Gr}(\mathbb{R}^2, 1)$ be

$$L_1 = \text{span} \left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}, \quad L_2 = \text{span} \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right\}.$$

For these subspaces and for $\|\cdot\| = \|\cdot\|_1$ it is easy to see that

$$\text{dist}_G(L_1, L_2) = 1 \neq \frac{1}{2} = \text{dist}_G(L_2, L_1)$$

and

$$\overline{\text{dist}}_G(L_1, L_2) = \frac{1}{2} \neq 1 = \overline{\text{dist}}_G(L_2, L_1).$$

The distances $\text{dist}_G(\cdot, \cdot)$ and $\overline{\text{dist}}_G(\cdot, \cdot)$ are bounded above as follows

$$\text{dist}_G(L_1, L_2) \leq \max_{\substack{x \in L_1 \\ x \neq 0}} \frac{\|x\|}{\|x\|}, \quad \overline{\text{dist}}_G(L_1, L_2) \leq \min_{\substack{v \in L_2 \\ v \neq 0}} \frac{\|v\|}{\|v\|}. \quad (5)$$

In the special case when $\|\cdot\| = \|\cdot\|_2$ the distances $\text{dist}_G(\cdot, \cdot)$ and $\overline{\text{dist}}_G(\cdot, \cdot)$ coincide and are both symmetric (see Proposition 3 below).

Define the set $\Sigma_m \subseteq \text{Gr}(V, m)$ of m -dimensional *ill-posed subspaces* with respect to the cone K as follows

$$\Sigma_m := \{L \in \text{Gr}(V, m) \mid L \cap K \neq \{0\} \text{ and } L^\perp \cap K^* \neq \{0\}\}.$$

Observe that $L \in \text{Gr}(V, m) \setminus \Sigma_m$ if and only if either $L \cap \text{int}(K) \neq \emptyset$ or $L^\perp \cap \text{int}(K^*) \neq \emptyset$. Define the *Grassmann distances to ill-posedness* of $L \in \text{Gr}(V, m)$ as follows

$$\text{dist}_G(L, \Sigma_m) := \min_{\tilde{L} \in \Sigma_m} \text{dist}_G(L, \tilde{L}), \quad \overline{\text{dist}}_G(\Sigma_m, L) := \min_{\tilde{L} \in \Sigma_m} \overline{\text{dist}}_G(\tilde{L}, L).$$

Observe that if $L \cap \text{int}(K) \neq \emptyset$ then

$$\text{dist}_G(L, \Sigma_m) = \min\{\text{dist}_G(L, \tilde{L}) : \tilde{L}^\perp \cap K^* \neq \{0\}\}.$$

Similarly, if $L^\perp \cap \text{int}(K^*) \neq \emptyset$ then

$$\overline{\text{dist}_G}(\Sigma_m, L) = \min\{\overline{\text{dist}_G}(\tilde{L}, L) : \tilde{L} \cap K \neq \{0\}\}.$$

The above distances to ill-posedness can alternatively be seen as *radii of well-posedness*: If $L \in \text{Gr}(V, m) \setminus \Sigma_m$ then for all $\tilde{L} \in \text{Gr}(V, m)$

$$\text{dist}_G(L, \tilde{L}) < \text{dist}_G(L, \Sigma_m) \Rightarrow \tilde{L} \notin \Sigma_m$$

and

$$\overline{\text{dist}_G}(\tilde{L}, L) < \overline{\text{dist}_G}(\Sigma_m, L) \Rightarrow \tilde{L} \notin \Sigma_m.$$

2.1 Least violated trial solutions

Define $\nu, \bar{\nu} : \text{Gr}(V, m) \rightarrow \mathbb{R}$ as follows

$$\nu(L) := \min_{\substack{u \in K^*, y \in L^\perp \\ \|u\|^* = 1}} \|y - u\|^*, \quad \bar{\nu}(L) := \min_{\substack{v \in K, x \in L \\ \|x\| = 1}} \|v - x\|.$$

Observe that $L \cap \text{int}(K) \neq \emptyset$ if and only if $\nu(L) > 0$. In this case $\nu(L)$ can be seen as a measure of a *least violated* trial point for the alternative problem $u \in L^\perp \cap K^* \setminus \{0\}$. Similarly, $L^\perp \cap \text{int}(K^*) \neq \emptyset$ if and only if $\bar{\nu}(L) > 0$. In this case $\bar{\nu}(L)$ can be seen as a measure of a *least violated* trial point for the alternative problem $x \in L \cap K \setminus \{0\}$. In Section 3.2 below we show that for a suitable norm $\|\cdot\|$ in V , $\nu(L)$ can also be seen as a measure of the *most interior* normalized solution to the feasibility problem $x \in L \cap \text{int}(K)$ when this problem is feasible.

The next result generalizes [13, Theorem 1] as well as [1, Proposition 1.6]:

Theorem 1. *Assume $L \in \text{Gr}(V, m)$.*

- (a) *If $L \cap \text{int}(K) \neq \emptyset$ then $\text{dist}_G(L, \Sigma_m) = \nu(L)$.*
- (b) *If $L^\perp \cap \text{int}(K^*) \neq \emptyset$ then $\overline{\text{dist}_G}(\Sigma_m, L) = \bar{\nu}(L)$.*

The proof of Theorem 1 relies on the properties of $\text{dist}_G(\cdot, \cdot)$ and $\nu(\cdot)$ stated in Proposition 1 and Proposition 2 below. These propositions in turn are consequences of the following technical lemma.

Lemma 1. Assume $L \in \text{Gr}(V, m)$ and $p \in V$. Then

$$\min_{z \in L} \|p - z\| = \max_{\substack{q \in L^\perp \\ \|q\|^* = 1}} \langle q, p \rangle. \quad (6)$$

Furthermore, $\bar{z} \in L$ and $\bar{q} \in L^\perp$ with $\|\bar{q}\|^* = 1$ attain the optimal values in (6) if and only if $\|p - \bar{z}\| = \langle \bar{q}, p - \bar{z} \rangle = \langle \bar{q}, p \rangle$.

Proof. By properties of norms and convex duality we have

$$\min_{z \in L} \|p - z\| = \min_{z \in L} \max_{\substack{q \in V \\ \|q\|^* \leq 1}} \langle q, p - z \rangle = \max_{\substack{q \in V \\ \|q\|^* \leq 1}} \min_{z \in L} \langle q, p - z \rangle = \max_{\substack{q \in L^\perp \\ \|q\|^* = 1}} \langle q, p \rangle.$$

This shows (6). The second statement follows from (6) and the fact that

$$\|p - \bar{z}\| \leq \langle \bar{q}, p - \bar{z} \rangle = \langle \bar{q}, p \rangle$$

for $\bar{z} \in L$ and $\bar{q} \in L^\perp$ with $\|\bar{q}\|^* = 1$. \square

Proposition 1. For $L_1, L_2 \in \text{Gr}(V, m)$

$$\text{dist}_G(L_1, L_2) = \max_{\substack{x \in L_1 \\ \|x\|=1}} \min_{v \in L_2} \|x - v\| = \max_{\substack{x \in L_1, u \in L_2^\perp \\ \|x\|=1, \|u\|^* = 1}} \langle u, x \rangle = \max_{\substack{u \in L_2^\perp \\ \|u\|^* = 1}} \min_{y \in L_1^\perp} \|u - y\|^*.$$

Furthermore, $\bar{x} \in L_1, \bar{v} \in L_2, \bar{u} \in L_2^\perp, \bar{y} \in L_1^\perp$ attain the above optimal values if and only if $\|\bar{x}\| = 1, \|\bar{u}\|^* = 1, \langle \bar{u}, \bar{x} - \bar{v} \rangle = \langle \bar{u}, \bar{x} \rangle = \|\bar{x} - \bar{v}\|$, and $\langle \bar{x}, \bar{u} - \bar{y} \rangle = \langle \bar{u}, \bar{x} \rangle = \|\bar{u} - \bar{y}\|^*$.

Proof. This follows by applying Lemma 1 to each of the norm minimization problems $\min_{v \in L_2} \|x - v\|$ and $\min_{y \in L_1^\perp} \|u - y\|^*$. \square

Proposition 2. If $L \in \text{Gr}(V, m)$ and $L \cap \text{int}(K) \neq \emptyset$ then

$$\nu(L) = \min_{\substack{u \in K^*, y \in L^\perp \\ \|u\|^* = 1}} \|u - y\|^* = \min_{\substack{u \in K^* \\ \|u\|^* = 1}} \max_{\substack{x \in L \\ \|x\|=1}} \langle u, x \rangle.$$

Furthermore, $\bar{u} \in K^*, \bar{y} \in L^\perp, \bar{x} \in L$ attain the above optimal values if and only if $\|\bar{u}\|^* = 1, \|\bar{x}\| = 1, \langle \bar{u} - \bar{y}, \bar{x} \rangle = \langle \bar{u}, \bar{x} \rangle = \|\bar{u} - \bar{y}\|^*$.

Proof. This follows by applying Lemma 1 to the norm minimization problem $\min_{y \in L^\perp} \|u - y\|^*$. \square

Proof of Theorem 1.

- (a) First we show $\text{dist}_G(L, \Sigma_m) \geq \nu(L)$. To do so, it suffices to show that $\text{dist}_G(L, \tilde{L}) \geq \nu(L)$ for all $\tilde{L} \in \text{Gr}(V, m)$ with $\tilde{L}^\perp \cap K^* \neq \{0\}$. Assume $\tilde{L} \in \text{Gr}(V, m)$ is such that $\tilde{L}^\perp \cap K^* \neq \{0\}$. Let $\bar{u} \in \tilde{L}^\perp \cap K^*$ with $\|\bar{u}\|^* = 1$. By Proposition 1, it follows that

$$\begin{aligned} \nu(L) &= \min_{\substack{u \in K^*, y \in L^\perp \\ \|\bar{u}\|^* = 1}} \|u - y\|^* \leq \min_{y \in L^\perp} \|\bar{u} - y\|^* \leq \max_{\substack{u \in \tilde{L}^\perp \\ \|\bar{u}\|^* = 1}} \min_{y \in L^\perp} \|u - y\|^* \\ &= \text{dist}_G(L, \tilde{L}). \end{aligned}$$

Next we show $\text{dist}_G(L, \Sigma_m) \leq \nu(L)$. To do so, it suffices to show that there exists $\tilde{L} \in \text{Gr}(V, m)$ such that $\tilde{L}^\perp \cap K^* \neq \{0\}$ and $\text{dist}_G(L, \tilde{L}) \leq \nu(L)$. For simplicity, we write ν as shorthand for $\nu(L)$ is the rest of this proof. By Proposition 2 there exist $\bar{x} \in L, \bar{u} \in K^*$, and $\bar{y} \in L^\perp$ such that $\|\bar{u}\|^* = 1, \|\bar{x}\| = 1$ and $\nu = \langle \bar{u} - \bar{y}, \bar{x} \rangle = \langle \bar{u}, \bar{x} \rangle = \|\bar{y} - \bar{u}\|^*$. Let $H := \{x \in V : \langle \bar{u} - \bar{y}, x \rangle = 0\}$. Observe that $\bar{x} \notin H$ because $\langle \bar{u} - \bar{y}, \bar{x} \rangle = \|\bar{y} - \bar{u}\|^* > 0$. Thus $L = \text{span}(\{\bar{x}\} \cup (L \cap H))$. Let $\bar{v} \in V$ be such that $\|\bar{v}\| = 1$ and $\langle \bar{u}, \bar{v} \rangle = \|\bar{u}\|^* = 1$. Next, define

$$\tilde{L} := \text{span}(\{\bar{x} - \nu\bar{v}\} \cup (L \cap H)).$$

Since $\bar{u} - \bar{y} \in H^\perp \subseteq (H \cap L)^\perp$ and $\bar{y} \in L^\perp \subseteq (H \cap L)^\perp$, we have $\bar{u} = \bar{u} - \bar{y} + \bar{y} \in (H \cap L)^\perp$; moreover, $\langle \bar{u}, \bar{x} - \nu\bar{v} \rangle = \langle \bar{u}, \bar{x} \rangle - \nu \langle \bar{u}, \bar{v} \rangle = 0$. Therefore $0 \neq \bar{u} \in \tilde{L}^\perp \cap K^*$. To upper bound $\text{dist}_G(L, \tilde{L})$, assume $x \in L \setminus \{0\}$. Then $x = \alpha\bar{x} + v$ for some $v \in H \cap L$. We thus have

$$\|x\| \geq \left\langle \frac{\bar{y} - \bar{u}}{\|\bar{y} - \bar{u}\|^*}, x \right\rangle = |\alpha| \frac{\langle \bar{y} - \bar{u}, \bar{x} \rangle}{\|\bar{y} - \bar{u}\|^*} = |\alpha|.$$

The second step follows because $\bar{y} - \bar{u} \in L^\perp$. Put $\tilde{x} := \alpha(\bar{x} - \nu\bar{v}) + v$ and observe that $\|\tilde{x} - x\| = |\alpha\nu|\|\bar{v}\| = |\alpha|\nu$. Hence

$$\frac{\|\tilde{x} - x\|}{\|x\|} = \frac{|\alpha|\nu}{\|x\|} \leq \nu.$$

Since this construction can be repeated for all $x \in L \setminus \{0\}$, it follows that

$$\text{dist}_G(L, \tilde{L}) = \max_{\substack{x \in L \\ x \neq 0}} \min_{\tilde{x} \in \tilde{L}} \frac{\|\tilde{x} - x\|}{\|x\|} \leq \nu.$$

- (b) This follows via a similar, though simpler, reasoning to that used in part (a). First, we show that $\overline{\text{dist}}_G(\Sigma_m, L) \geq \bar{\nu}(L)$. To do so, it suffices to show that $\text{dist}_G(\tilde{L}, L) \geq \bar{\nu}(L)$ for all $\tilde{L} \in \text{Gr}(V, m)$ with $\tilde{L} \cap K \neq \{0\}$. Assume $\tilde{L} \in \text{Gr}(V, m)$ is such that $\tilde{L} \cap K \neq \{0\}$. Let $\bar{v} \in \tilde{L} \cap K$ with $\bar{v} \neq 0$. Thus

$$\begin{aligned} \bar{\nu}(L) &= \min_{\substack{v \in K, x \in L \\ \|x\|=1}} \|x - v\| \leq \min_{\substack{x \in L \\ x \neq 0}} \frac{\|x - \bar{v}\|}{\|x\|} \leq \max_{\substack{v \in \tilde{L} \\ v \neq 0}} \inf_{\substack{x \in L \\ x \neq 0}} \frac{\|x - v\|}{\|x\|} \\ &= \overline{\text{dist}}_G(\tilde{L}, L). \end{aligned}$$

Next we show $\overline{\text{dist}}_G(\Sigma_m, L) \leq \bar{\nu}(L)$. To do so, it suffices to show that there exists $\tilde{L} \in \text{Gr}(V, m)$ such that $\tilde{L} \cap K \neq \{0\}$ and $\overline{\text{dist}}_G(\tilde{L}, L) \leq \bar{\nu}(L)$. For simplicity, we write $\bar{\nu}$ as shorthand for $\bar{\nu}(L)$ in the rest of this proof. Let $\bar{x} \in L$ and $\bar{v} \in K$ be such that $\|\bar{x}\| = 1$ and $\bar{\nu} = \|\bar{x} - \bar{v}\|^*$. If $\bar{v} = 0$ then from (5) it readily follows that $\bar{\nu} = \min_{x \in L, \|x\|=1} \|x\| \geq \overline{\text{dist}}_G(\tilde{L}, L)$ for any $\tilde{L} \in \Sigma_m$ and there is nothing to show. Thus assume $\bar{v} \neq 0$. Let $\bar{y} \in V$ be such that $\|\bar{y}\|^* = 1$ and $\langle \bar{y}, \bar{x} \rangle = \|\bar{x}\| = 1$. Let $H := \{x \in V : \langle \bar{y}, x \rangle = 0\}$. Observe that $\bar{x} \notin H$ because $\langle \bar{y}, \bar{x} \rangle = 1$. Thus $L = \text{span}(\{\bar{x}\} \cup (L \cap H))$. Next, define

$$\tilde{L} := \text{span}(\{\bar{v}\} \cup (L \cap H)).$$

By construction, we have $0 \neq \bar{v} \in \tilde{L} \cap K$. To upper bound $\overline{\text{dist}}_G(\tilde{L}, L)$, assume $\tilde{x} \in \tilde{L} \setminus \{0\}$. Then $\tilde{x} = \alpha \bar{v} + v$ for some $v \in H \cap L$. Put $x := \alpha \bar{x} + v$ and observe that $\|x\| \geq |\langle \bar{y}, x \rangle| = |\alpha|$ since $\|\bar{y}\|^* = 1$ and $\langle \bar{y}, v \rangle = 0$. Therefore when $\alpha \neq 0$,

$$\frac{\|\tilde{x} - x\|}{\|x\|} = \frac{|\alpha| \|\bar{x} - \bar{v}\|}{\|x\|} \leq \frac{|\alpha| \bar{\nu}}{|\alpha|} \leq \bar{\nu}.$$

On the other hand, when $\alpha = 0$

$$\frac{\|\tilde{x} - x\|}{\|x\|} = \frac{0}{\|v\|} = 0 \leq \bar{\nu}.$$

Since this construction can be repeated for all $\tilde{x} \in \tilde{L} \setminus \{0\}$, it follows that

$$\overline{\text{dist}}_G(\tilde{L}, L) = \max_{\substack{\tilde{x} \in \tilde{L} \\ \tilde{x} \neq 0}} \inf_{\substack{x \in L \\ x \neq 0}} \frac{\|\tilde{x} - x\|}{\|x\|} \leq \bar{\nu}.$$

□

2.2 Renegar's distance to ill-posedness

We now relate the condition measures $\text{dist}_G(\cdot), \overline{\text{dist}}_G(\cdot)$ with the now classical Renegar's distance to ill-posedness. The key conceptual difference between Renegar's approach and the Grassmann approach is that Renegar [25, 26] considers conic feasibility problems where the linear spaces L and L^\perp are defined via a linear mapping $A : W \rightarrow V$. We next present a variation (more precisely a slight extension) of Renegar's distance to ill-posedness.

Assume W is a finite dimensional Euclidean vector space with norm $|\cdot|$ and $\dim(W) < \dim(V)$. For a linear mapping $A : W \rightarrow V$ consider the conic systems (2) and (3) defined by taking $L = \text{Im}(A)$. These two conic systems can respectively be written as

$$Ax \in K \setminus \{0\} \quad (7)$$

and

$$A^*u = 0, \quad u \in K^* \setminus \{0\}. \quad (8)$$

Here $A^* : V \rightarrow W$ denotes the *adjoint* operator of A , that is, the linear mapping satisfying $\langle y, Aw \rangle = \langle A^*y, w \rangle$ for all $y \in V, w \in W$.

Let $\mathcal{L}(W, V)$ denote the set of linear mappings from W to V . Define the operator norms $\|\cdot\|, \|\cdot\| : \mathcal{L}(W, V) \rightarrow \mathbb{R}$ as follows

$$\|A\| := \max_{\substack{w \in W \\ |w|=1}} \|Aw\| \quad \text{and} \quad \|A\| := \max_{\substack{w \in W \\ |w|=1}} \|Aw\|.$$

We say that $A \in \mathcal{L}(W, V)$ is *well-posed* if either (7) or (8) is feasible and remains so for sufficiently small perturbations of A . Let $\Sigma \subseteq \mathcal{L}(W, V)$ denote the set of *ill-posed mappings*, that is, mappings that are not well-posed. It is easy to see that if $A \in \mathcal{L}(W, V) \setminus \Sigma$ then either $\text{Im}(A) \cap \text{int}(K) \neq \emptyset$ or $\ker(A^*) \cap \text{int}(K^*) \neq \emptyset$.

Renegar's distance to ill-posedness $\text{dist}_R : \mathcal{L}(W, V) \rightarrow \mathbb{R}$ is defined as:

$$\text{dist}_R(A, \Sigma) := \min \left\{ \|A - \tilde{A}\| : \tilde{A} \in \Sigma \right\}.$$

This definition is a slight extension of Renegar's original definition in [25, 26] due to our use of the two norms $\|\cdot\|, \|\cdot\|$ in V . Renegar's original distance to ill-posedness corresponds to the case $\|\cdot\| = \|\cdot\|$.

Renegar's distance to ill-posedness $\text{dist}_R(A, \Sigma)$ can alternatively be defined as the *distance to infeasibility* of (7) or (8). More precisely, if $A \in \mathcal{L}(W, V)$ is such that (7) is feasible, then

$$\text{dist}_R(A, \Sigma) := \min \left\{ \|A - \tilde{A}\| : \tilde{A}x \in K \setminus \{0\} \text{ is infeasible} \right\}.$$

Likewise if $A \in \mathcal{L}(W, V)$ is such that (8) is feasible, then

$$\text{dist}_R(A, \Sigma) := \min \left\{ \|A - \tilde{A}\| : \tilde{A}^* u = 0, u \in K^* \setminus \{0\} \text{ is infeasible} \right\}.$$

Observe that if $A \in \mathcal{L}(W, V) \setminus \Sigma$ and $\text{Im}(A) \cap \text{int}(K) \neq \emptyset$ then

$$\text{dist}_R(A, \Sigma) = \min \left\{ \|A - \tilde{A}\| : \tilde{A}^* u = 0 \text{ for some } u \in K^* \setminus \{0\} \right\}.$$

On the other hand, if $A \in \mathcal{L}(W, V) \setminus \Sigma$ and $\text{Im}(A)^\perp \cap \text{int}(K^*) \neq \emptyset$ then

$$\text{dist}_R(A, \Sigma) = \min \left\{ \|A - \tilde{A}\| : \tilde{A}x \in K \text{ for some } x \in W \setminus \{0\} \right\}.$$

We shall rely on one more piece of notation. Given $A \in \mathcal{L}(W, V)$ the mapping $A^{-1} : \text{Im}(A) \rightarrow W$ is well defined provided A is injective. In this case let $\|A^{-1}\|$ denote the following operator norm

$$\|A^{-1}\| := \max_{\substack{x \in \text{Im}(A) \\ \|x\|=1}} |A^{-1}x|.$$

The following result gives a general version of [1, Theorem 1.4]. We note that the statements of these theorems are nearly identical.

Theorem 2. *Assume $A \in \mathcal{L}(W, V) \setminus \Sigma$ is injective and $L := \text{Im}(A) \in \text{Gr}(V, m)$.*

(a) *If $L \cap \text{int}(K) \neq \emptyset$ then*

$$\frac{1}{\|A\|} \leq \frac{\text{dist}_G(L, \Sigma_m)}{\text{dist}_R(A, \Sigma)} \leq \|A^{-1}\|,$$

or equivalently

$$\frac{1}{\text{dist}_G(L, \Sigma_m)} \leq \frac{\|A\|}{\text{dist}_R(A, \Sigma)} \leq \frac{\kappa(A)}{\text{dist}_G(L, \Sigma_m)}$$

where $\kappa(A) = \|A\| \cdot \|A^{-1}\|$.

(b) *If $L^\perp \cap \text{int}(K^*) \neq \emptyset$ then*

$$\frac{1}{\|A\|} \leq \frac{\overline{\text{dist}_G(\Sigma_m, L)}}{\text{dist}_R(A, \Sigma)} \leq \|A^{-1}\|,$$

or equivalently

$$\frac{1}{\overline{\text{dist}_G(\Sigma_m, L)}} \leq \frac{\|A\|}{\text{dist}_R(A, \Sigma)} \leq \frac{\kappa(A)}{\overline{\text{dist}_G(\Sigma_m, L)}}.$$

Proof. **(a)** First, we prove $\text{dist}_R(A, \Sigma) \leq \nu(L)\|A\| = \text{dist}_G(L, \Sigma_m)\|A\|$. To that end, assume $\bar{u} \in K^*$, $\|\bar{u}\|^* = 1$ is such that $\nu(L) = \max_{\substack{x \in L \\ \|x\|=1}} \langle \bar{u}, x \rangle$ as in Proposition 2. Then

$$\nu(L)\|A\| \geq \max_{\substack{w \in W \\ \|w\|=1}} \langle \bar{u}, Aw \rangle = |A^* \bar{u}|^*.$$

Let $\bar{v} \in V$, $\|\bar{v}\| = 1$ be such that $\langle \bar{u}, \bar{v} \rangle = \|\bar{u}\|^* = 1$. Now construct $\Delta A : W \rightarrow V$ as follows

$$\Delta A(w) := -\langle A^* \bar{u}, w \rangle \bar{v}.$$

Observe that $\|\Delta A\| = |A^* \bar{u}|^* \cdot \|\bar{v}\| \leq \nu(L)\|A\|$, and $\Delta A^* : V \rightarrow W$ is defined by

$$\Delta A^*(y) = -\langle y, \bar{v} \rangle A^* \bar{u}.$$

In particular $(A + \Delta A)^* \bar{u} = A^* \bar{u} - \langle \bar{u}, \bar{v} \rangle A^* \bar{u} = 0$ and $\bar{u} \in K^* \setminus \{0\}$. Therefore

$$\text{dist}_R(A, \Sigma_m) \leq \|\Delta A\| \leq \nu(L)\|A\| = \text{dist}_G(L, \Sigma_m)\|A\|.$$

For the second inequality, assume $\tilde{A} \in \mathcal{L}(W, V)$ is such that $\ker(\tilde{A}^*) \cap K^* \setminus \{0\} \neq \emptyset$. By performing a slight perturbation on \tilde{A} if needed, we may assume that $\tilde{L} \in \text{Gr}(V, m)$. Thus

$$\text{dist}_G(L, \Sigma_m) \leq \text{dist}_G(L, \tilde{L}).$$

Assume $\bar{x} \in L$ with $\|\bar{x}\| = 1$ is such that

$$\text{dist}_G(L, \tilde{L}) = \min_{\tilde{x} \in \tilde{L}} \|\bar{x} - \tilde{x}\|.$$

Let $\bar{w} \in W$ be such that $A\bar{w} = \bar{x}$. Since $\tilde{A}\bar{w} \in \tilde{L}$, it follows that

$$\text{dist}_G(L, \tilde{L}) \leq \|(A - \tilde{A})\bar{w}\| \leq \|(A - \tilde{A})\| \cdot \|\bar{w}\|.$$

Thus

$$\|A - \tilde{A}\| \geq \frac{\text{dist}_G(L, \tilde{L})}{\|\bar{w}\|} \geq \frac{\text{dist}_G(L, \Sigma_m)}{\|A^{-1}\|}.$$

Since this holds for all \tilde{A} with $\ker(\tilde{A}^*) \cap K^* \setminus \{0\}$, we conclude that

$$\text{dist}_R(A, \Sigma_m) \geq \frac{\text{dist}_G(L, \Sigma_m)}{\|A^{-1}\|}.$$

(b) First, we prove $\text{dist}_R(A, \Sigma_m) \leq \bar{\nu}(L)\|A\| = \overline{\text{dist}_G(\Sigma_m, L)}\|A\|$. To that end, assume $\bar{v} \in K, \bar{x} \in L$ are such that $\|\bar{x}\| = 1$ and $\bar{\nu} = \|\bar{x} - \bar{v}\|$.

Let $\bar{w} := A^{-1}\bar{x}$ and $\bar{z} \in W$ be such that $|\bar{z}|^* = 1$ and $\langle \bar{z}, \bar{w} \rangle = |\bar{w}|$. Observe that $|\bar{w}| \geq \frac{1}{\|A\|}$ since $1 = \|x\| = \|A\bar{w}\| \leq \|A\||\bar{w}|$. Define $\Delta A : W \rightarrow V$ as follows

$$\Delta A(w) := (\bar{v} - \bar{x}) \frac{\langle \bar{z}, w \rangle}{|\bar{w}|}.$$

Observe that $(A + \Delta A)\bar{w} = \bar{v} \in K$ and $\|\Delta A\| = \frac{\|\bar{v} - \bar{x}\|}{|\bar{w}|} \leq \bar{\nu}\|A\|$. Thus $\text{dist}_R(A, \Sigma_m) \leq \|\Delta A\| \leq \bar{\nu}\|A\| = \overline{\text{dist}_G(\Sigma_m, L)}\|A\|$. For the second inequality assume $\tilde{A} \in \mathcal{L}(W, V)$ is such that $\text{Im}(\tilde{A}) \cap K \setminus \{0\} \neq \emptyset$. By performing a slight perturbation of \tilde{A} if needed, we may assume that $\tilde{L} := \text{Im}(\tilde{A}) \in \text{Gr}(V, m)$. Thus

$$\overline{\text{dist}_G(\Sigma_m, L)} \leq \overline{\text{dist}_G(\tilde{L}, L)}.$$

Assume $0 \neq \tilde{x} \in \tilde{L}$ is such that

$$\overline{\text{dist}_G(\tilde{L}, L)} = \inf_{x \in L} \frac{\|\tilde{x} - x\|}{\|x\|}.$$

Let $w \in W$ be such that $\tilde{A}w = \tilde{x}$. Since $Aw \in L$ we have

$$\overline{\text{dist}_G(\tilde{L}, L)} \leq \frac{\|\tilde{A}w - Aw\|}{\|Aw\|} \leq \frac{\|\tilde{A} - A\||w|}{\|Aw\|}.$$

Hence

$$\|\tilde{A} - A\| \geq \frac{\overline{\text{dist}_G(\tilde{L}, L)}\|Aw\|}{|w|} = \frac{\overline{\text{dist}_G(\tilde{L}, L)}\|Aw\|}{|A^{-1}(Aw)|} \geq \frac{\overline{\text{dist}_G(\Sigma_m, L)}}{\|A^{-1}\|}.$$

Since this holds for all \tilde{A} with $\text{Im}(\tilde{A}) \cap K \setminus \{0\} \neq \emptyset$, we conclude that

$$\text{dist}_R(A, \Sigma_m) \geq \frac{\overline{\text{dist}_G(\Sigma_m, L)}}{\|A^{-1}\|}.$$

□

3 Euclidean and induced norms

We next discuss the special but important choices of norms, namely the Euclidean norm defined by the inner product in V and a certain *induced norm* defined by the cone K and a point $e \in \text{int}(K)$.

3.1 Euclidean norm

The Euclidean norm $\|\cdot\|_2$ is the norm defined by the inner product in V . That is, for $v \in V$

$$\|v\|_2 := \sqrt{\langle v, v \rangle}.$$

The Euclidean norm is naturally related to angles. Given $x, y \in V \setminus \{0\}$ define $\angle(x, y) := \arccos \frac{\langle x, y \rangle}{\|x\|_2 \|y\|_2} \in [0, \pi]$. Given a linear subspace $L \subseteq V$ and a closed convex cone $C \subseteq V$ define

$$\angle(L, C) := \min\{\angle(x, v) : x \in L \setminus \{0\}, v \in C \setminus \{0\}\} \in [0, \pi/2].$$

Proposition 3. *Assume $\|\cdot\| = \|\cdot\|_2$. Then*

(a) *For $L_1, L_2 \in \text{Gr}(V, m)$*

$$\|\Pi_{L_1} - \Pi_{L_2}\|_2 = \text{dist}_G(L_1, L_2) = \overline{\text{dist}_G}(L_1, L_2)$$

where $\Pi_{L_i} : V \rightarrow L_i$ is the orthogonal projection onto L_i for $i = 1, 2$ and $\|\Pi_{L_1} - \Pi_{L_2}\|_2 = \max_{\|x\|_2=1} \|(\Pi_{L_1} - \Pi_{L_2})x\|_2$.

(b) *If $L \in \text{Gr}(V, m)$ is such that $L \cap \text{int}(K) \neq \emptyset$ then*

$$\nu(L) = \sin \angle(L^\perp, K^*).$$

(c) *If $L \in \text{Gr}(V, m)$ is such that $L^\perp \cap \text{int}(K^*) \neq \emptyset$ then*

$$\bar{\nu}(L) = \sin \angle(L, K).$$

Proof. (a) A classical result from geometry (see, e.g., [21, Section 12.4]) states that $\|\Pi_{L_1} - \Pi_{L_2}\|_2 = \sin(\theta)$ where θ is the *largest principal angle* between L_1 and L_2 , namely

$$\theta = \max_{\substack{x \in L_1 \\ x \neq 0}} \min_{\substack{v \in L_2 \\ v \neq 0}} \angle(x, v) \in [0, \pi/2].$$

Since $\theta \in [0, \pi/2]$ we have

$$\begin{aligned} \|\Pi_{L_1} - \Pi_{L_2}\|_2 &= \max_{\substack{x \in L_1 \\ x \neq 0}} \min_{\substack{v \in L_2 \\ v \neq 0}} \sin(\angle(x, v)) \\ &= \max_{\substack{x \in L_1 \\ x \neq 0}} \min_{v \in L_2} \frac{\|x - v\|_2}{\|x\|_2} \\ &= \text{dist}_G(L_1, L_2), \end{aligned}$$

and

$$\begin{aligned}
\|\Pi_{L_1} - \Pi_{L_2}\|_2 &= \max_{\substack{x \in L_1 \\ x \neq 0}} \min_{\substack{v \in L_2 \\ v \neq 0}} \sin(\angle(x, v)) \\
&= \max_{\substack{x \in L_1 \\ x \neq 0}} \inf_{\substack{v \in L_2 \\ v \neq 0}} \frac{\|x - v\|_2}{\|v\|_2} \\
&= \overline{\text{dist}_G}(L_1, L_2).
\end{aligned}$$

(b) Since $\angle(L^\perp, K^*) \in [0, \pi/2]$ we have

$$\sin \angle(L^\perp, K^*) = \min_{\substack{u \in K^*, y \in L^\perp \\ u, y \neq 0}} \sin \angle(y, u) = \min_{\substack{u \in K^*, y \in L^\perp \\ \|u\|_2=1}} \|u - y\|_2 = \nu(L).$$

(c) This is nearly identical to part (b): Since $\angle(L, K) \in [0, \pi/2]$ we have

$$\sin \angle(L, K) = \min_{\substack{v \in K, x \in L \\ v, x \neq 0}} \sin \angle(x, v) = \min_{\substack{v \in K, x \in L \\ \|x\|_2=1}} \|v - x\|_2 = \bar{\nu}(L).$$

□

Proposition 3 readily implies the symmetry of dist_G and $\overline{\text{dist}_G}$. That is, $\text{dist}_G(L_1, L_2) = \text{dist}_G(L_2, L_1)$ for all $L_1, L_2 \in \text{Gr}(V, m)$, and likewise for $\overline{\text{dist}_G}(\cdot)$ if $\|\cdot\| = \|\cdot\|_2$. Furthermore, Proposition 3 also implies that in this case $\text{dist}_G(\cdot)$ coincides with the Grassmann distance used in [1]. Hence for this special choice of norms, Theorem 2 subsumes [1, Theorem 1.4]. Similarly, for this special choice of norms, Theorem 1 and Proposition 3 subsume [1, Proposition 1.6].

3.2 Induced norm and induced eigenvalue mappings

Assume $K \subseteq V$ is a regular closed convex cone and $e \in \text{int}(K)$. We next describe a norm $\|\cdot\|_e$ in V and a mapping $\lambda_e : V \rightarrow \mathbb{R}$ induced by the pair (K, e) . These norm and mapping yield a natural interpretation of $\nu(L)$ as a measure of the *most interior* normalized solution to the feasibility problem $x \in L \cap \text{int}(K)$ when this problem is feasible.

Define the norm $\|\cdot\|_e$ in V induced by (K, e) as follows (see [12])

$$\|x\|_e := \min\{\alpha \geq 0 : x + \alpha e \in K, -x + \alpha e \in K\}.$$

Define the *eigenvalue* mapping $\lambda_e : V \rightarrow \mathbb{R}$ induced by (K, e) as follows

$$\lambda_e(x) := \max\{t \in \mathbb{R} : x - te \in K\}.$$

Observe that $x \in K \Leftrightarrow \lambda_e(x) \geq 0$ and $x \in \text{int}(K) \Leftrightarrow \lambda_e(x) > 0$. Furthermore, observe that when $x \in K$

$$\lambda_e(x) = \max\{r \geq 0 : \|v\|_e \leq r \Rightarrow x + v \in K\}.$$

Thus for $x \in K$, $\lambda_e(x)$ is a measure of how interior x is in the cone K .

It is easy to see that $\|u\|_e^* = \langle u, e \rangle$ for $u \in K^*$. In analogy to the standard simplex, let

$$\Delta(K^*) := \{u \in K^* : \|u\|_e^* = 1\} = \{u \in K^* : \langle u, e \rangle = 1\}.$$

It is also easy to see that the eigenvalue mapping λ_e has the following alternative expression

$$\lambda_e(x) = \min_{u \in \Delta(K^*)} \langle u, x \rangle.$$

The following result readily follows from Proposition 2 and convex duality.

Proposition 4. *Let $K \subseteq V$ be a regular closed convex cone and $e \in \text{int}(K)$. If $\|\cdot\| = \|\cdot\|_e$, then for all $L \in \text{Gr}(V, m)$*

$$\nu(L) = \min_{u \in \Delta(K^*)} \max_{\substack{x \in L \\ \|x\| \leq 1}} \langle x, u \rangle = \max_{\substack{x \in L \\ \|x\| \leq 1}} \min_{u \in \Delta(K^*)} \langle x, u \rangle = \max_{\substack{x \in L \\ \|x\| \leq 1}} \lambda_e(x).$$

In particular, when $L \cap \text{int}(K) \neq \emptyset$ the quantity $\nu(L)$ can be seen as a measure of the most interior normalized point in $L \cap \text{int}(K)$.

It is illustrative to further specialize Proposition 4 to two important cases. The first case is $V = \mathbb{R}^n$ with the usual dot inner product, $K = \mathbb{R}_+^n$ and $e = [1 \ \dots \ 1]^\top \in \mathbb{R}_+^n$. In this case $\|\cdot\|_e = \|\cdot\|_\infty$, $\|\cdot\|_e^* = \|\cdot\|_1$, $(\mathbb{R}_+^n)^* = \mathbb{R}_+^n$ and $\Delta(\mathbb{R}_+^n)$ is the standard simplex $\Delta_{n-1} := \{x \in \mathbb{R}_+^n : \sum_{i=1}^n x_i = 1\}$. In this case $\lambda_e(x) = \min_{i=1, \dots, n} x_i$. Hence for $\|\cdot\| = \|\cdot\|_e$ we have

$$\nu(L) = \max_{\substack{x \in L \\ \|x\| \leq 1}} \min_{j=1, \dots, n} x_j. \quad (9)$$

The second special case is $V = \mathbb{S}^n$ with the trace inner product, $K = \mathbb{S}_+^n$ and $e = I \in \mathbb{S}_+^n$. In this case $\|\cdot\|_e$ and $\|\cdot\|_e^*$ are respectively the operator norm and the nuclear norm in \mathbb{S}^n . More precisely

$$\|X\|_e = \max_{i=1, \dots, n} |\lambda_i(X)|, \quad \|X\|_e^* = \sum_{i=1}^n |\lambda_i(X)|,$$

where $\lambda_i(X)$, $i = 1, \dots, n$ are the usual eigenvalues of X . Furthermore, $(\mathbb{S}_+^n)^* = \mathbb{S}_+^n$ and $\Delta(\mathbb{S}_+^n)$ is the *spectraplex* $\{X \in \mathbb{S}_+^n : \sum_{i=1}^n \lambda_i(X) = 1\}$.

In this case $\lambda_e(x) = \min_{j=1,\dots,n} \lambda_j(X)$. Thus, in nice analogy to (9), for $\|\cdot\| = \|\cdot\|_e$ we have

$$\nu(L) = \max_{\substack{X \in L \\ \|X\| \leq 1}} \min_{j=1,\dots,n} \lambda_j(X). \quad (10)$$

Section 5.1 below details how (9) and (10) extend to the more general case when K is a symmetric cone.

3.3 Sigma measure

The induced eigenvalue function discussed in Section 3.2 above can be defined more broadly. Given $v \in K \setminus \{0\}$ define $\lambda_v : V \rightarrow [-\infty, \infty)$ as follows

$$\lambda_v(x) := \max\{t : x - tv \in K\}.$$

Define $\sigma : \text{Gr}(V, m) \rightarrow [0, \infty)$ as follows

$$\sigma(L) := \min_{\substack{v \in K \\ \|v\|=1}} \max_{\substack{x \in L \\ \|x\| \leq 1}} \lambda_v(x). \quad (11)$$

As we show in Proposition 8 below, $\sigma(L)$ can be seen as a generalization of the sigma measure introduced by Ye [29]. Observe that $L \cap \text{int}(K) \neq \emptyset$ if and only if $\sigma(L) > 0$ and in this case

$$\sigma(L) = \min_{\substack{v \in K \\ \|v\|=1}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \lambda_v(x).$$

As Corollary 1 below shows, when $L \cap \text{int}(K) \neq \emptyset$, the quantities $\sigma(L)$ and $\nu(L)$ are closely related. To that end, we rely on the following analogous result to Proposition 2.

Proposition 5. *If $L \in \text{Gr}(V, m)$ and $L \cap \text{int}(K) \neq \emptyset$ then*

$$\sigma(L) = \min_{\substack{v \in K \\ \|v\|=1}} \max_{\substack{x \in L \\ \|x\|=1}} \lambda_v(x) = \min_{\substack{v \in K, y \in L^\perp, u \in K^* \\ \|v\|=1, \langle u, v \rangle = 1}} \|u - y\|^*. \quad (12)$$

Proof. Assume $v \in K$ is fixed. By the construction of λ_v , convex duality,

and Lemma 1 we have

$$\begin{aligned}
\max_{\substack{x \in L \\ \|x\|=1}} \lambda_v(x) &= \max_{\substack{x \in L, t \in \mathbb{R} \\ \|x\| \leq 1, x - tv \in K}} t \\
&= \max_{\substack{x \in L, t \in \mathbb{R} \\ \|x\| \leq 1}} \min_{u \in K^*} (t + \langle u, x - tv \rangle) \\
&= \min_{u \in K^*} \max_{\substack{x \in L, t \in \mathbb{R} \\ \|x\| \leq 1}} (t + \langle u, x - tv \rangle) \\
&= \min_{\substack{u \in K^* \\ \langle u, v \rangle = 1}} \max_{\substack{x \in L \\ \|x\|=1}} \langle u, x \rangle \\
&= \min_{\substack{u \in K^*, y \in L^\perp \\ \langle u, v \rangle = 1}} \|u - y\|^*.
\end{aligned}$$

We thus get (12) by taking minimum over the set $\{v \in K : \|v\| = 1\}$. \square

Corollary 1. Assume $L \in \text{Gr}(V, m)$ is such that $L \cap \text{int}(K) \neq \emptyset$.

(a) For any two norms $\|\cdot\|, \|\cdot\|$ in V the following holds

$$1 \leq \min_{\substack{v \in K, u \in K^* \\ \|v\|=1, \langle u, v \rangle = 1}} \|u\|^* \leq \frac{\sigma(L)}{\nu(L)} \leq \frac{1}{\min_{\substack{u \in K^* \\ \|u\|=1}} \max_{\substack{v \in K \\ \|v\|=1}} \langle u, v \rangle}.$$

(b) If $\|\cdot\| = \|\cdot\| = \|\cdot\|_2$ then

$$1 \leq \frac{\sigma(L)}{\nu(L)} \leq \frac{1}{\cos(\Theta(K^*, K))}.$$

where

$$\Theta(K^*, K) := \max_{u \in K^* \setminus \{0\}} \min_{v \in K \setminus \{0\}} \angle(u, v).$$

In particular, if in addition $K = K^*$ then $\nu(L) = \sigma(L)$.

(c) If $\|\cdot\| = \|\cdot\|_e$ then

$$\sigma(L) = \nu(L).$$

Proof. (a) The first inequality is an immediate consequence of Hölder's inequality (4). Next, from Proposition 5 it follows that $\sigma(L) = \|\bar{u} - \bar{y}\|^*$ for some $\bar{v} \in K, \bar{y} \in L^\perp, \bar{u} \in K^*$ with $\|\bar{v}\| = 1, \langle \bar{u}, \bar{v} \rangle = 1$. Thus from the construction of $\nu(L)$ we get

$$\nu(L) \leq \frac{\|\bar{u} - \bar{y}\|^*}{\|\bar{u}\|^*} \leq \frac{\sigma(L)}{\min_{\substack{v \in K, u \in K^* \\ \|v\|=1, \langle u, v \rangle = 1}} \|u\|^*}$$

and hence the second inequality follows.

For the third inequality assume $\nu(L) = \|\hat{u} - \hat{y}\|^*$ for some $\hat{u} \in K^*$, $\hat{y} \in L^\perp$ with $\|\hat{u}\|^* = 1$. Then by Proposition 5 we get

$$\begin{aligned}
\sigma(L) &= \min_{\substack{v \in K, y \in L^\perp, u \in K^* \\ \|v\|=1, \langle u, v \rangle = 1}} \|u - y\|^* \leq \inf_{\substack{v \in K, y \in L^\perp, \\ \|v\|=1, \langle \hat{u}, v \rangle \neq 0}} \left\| \frac{\hat{u}}{\langle \hat{u}, v \rangle} - y \right\|^* \\
&= \inf_{\substack{v \in K, y \in L^\perp, \\ \|v\|=1, \langle \hat{u}, v \rangle = 1}} \frac{\|\hat{u} - y\|^*}{\langle \hat{u}, v \rangle} = \frac{\min_{y \in L^\perp} \|\hat{u} - y\|^*}{\max_{\substack{v \in K \\ \|v\|=1}} \langle \hat{u}, v \rangle} \\
&\leq \frac{\|\hat{u} - \hat{y}\|^*}{\max_{\substack{v \in K \\ \|v\|=1}} \langle \hat{u}, v \rangle},
\end{aligned}$$

hence

$$\sigma(L) \leq \frac{\|\hat{u} - \hat{y}\|^*}{\max_{\substack{v \in K \\ \|v\|=1}} \langle \hat{u}, v \rangle} \leq \frac{\nu(L)}{\min_{\substack{u \in K^* \\ \|u\|^*=1}} \max_{\substack{v \in K \\ \|v\|=1}} \langle u, v \rangle}$$

and the third inequality follows.

- (b) The first inequality follows from part (a). For the second inequality observe that since $\cos(\cdot)$ is decreasing in $[0, \pi]$

$$\begin{aligned}
\cos(\Theta(K^*, K)) &= \min_{u \in K^* \setminus \{0\}} \max_{v \in K \setminus \{0\}} \cos(\angle(u, v)) \\
&= \min_{u \in K^* \setminus \{0\}} \max_{v \in K \setminus \{0\}} \frac{\langle u, v \rangle}{\|u\|_2 \cdot \|v\|_2} \\
&= \min_{\substack{u \in K^* \\ \|u\|_2=1}} \max_{\substack{v \in K \\ \|v\|_2=1}} \langle u, v \rangle.
\end{aligned}$$

The second inequality then follows from part (a) as well.

If in addition $K = K^*$ then $\Theta(K^*, K) = 0$ and consequently $\frac{\sigma(L)}{\nu(L)} = 1$.

- (c) Since $\|\cdot\| = \|\cdot\|_e$, we have $\|e\| = 1$ and $\|u\|^* = \langle u, e \rangle$ for all $u \in K^*$. Thus $\min_{\substack{u \in K^* \\ \|u\|^*=1}} \max_{\substack{v \in K \\ \|v\|=1}} \langle u, v \rangle \geq \min_{\substack{u \in K^* \\ \|u\|^*=1}} \langle u, e \rangle = 1$. Therefore from part (a) it follows that $\frac{\sigma(L)}{\nu(L)} = 1$. □

The following example shows that the upper bound in Corollary 1(b) is tight.

Example 1. Assume $\phi \in (0, \pi/4)$. Let $V = \mathbb{R}^2$ be endowed with the dot inner product and $K := \{(x_1, x_2) \in V : \cos(\phi)x_2 \geq |x_1|\}$. Then $K^* = \{(x_1, x_2) \in V : x_2 \geq \cos(\phi)|x_1|\}$. In this case $\Theta(K, K^*) = \pi/2 - 2\phi$. For $L = \{x \in V : x_1 = 0\}$ and $\|\cdot\| = \|\cdot\| = \|\cdot\|_2$ it is easy to see that $\nu(L) = \sin(\phi)$ and $\sigma(L) = 1/(2\cos(\phi))$. Hence

$$\frac{\sigma(L)}{\nu(L)} = \frac{1}{2\sin(\phi)\cos(\phi)} = \frac{1}{\sin(2\phi)} = \frac{1}{\cos(\pi/2 - 2\phi)} = \frac{1}{\cos(\Theta(K, K^*))}.$$

4 Symmetry measure

Next, we will consider the symmetry measure of L , which has been used as a measure of conditioning [5, 6]. This measure is defined as follows. Given a set S in a vector space such that $0 \in S$, define

$$\text{Sym}(0, S) := \max\{t \geq 0 : w \in S \Rightarrow -tw \in S\}. \quad (13)$$

Observe that $\text{Sym}(0, S) \in [0, 1]$ with $\text{Sym}(0, S) = 1$ precisely when S is perfectly symmetric around 0.

Assume $A : V \rightarrow W$ is a linear mapping such that $L = \ker(A)$. Define the *symmetry* measure of L relative to K as follows.

$$\text{Sym}(L) := \text{Sym}(0, A(\{x \in K : \|x\| \leq 1\})). \quad (14)$$

It is easy to see that $\text{Sym}(L)$ depends only on L, K and not on the choice of A . More precisely, $\text{Sym}(0, A(\{x \in K : \|x\| \leq 1\})) = \text{Sym}(0, A'(\{x \in K : \|x\| \leq 1\}))$ if $\ker(A) = \ker(A') = L$. Indeed, the quantity $\text{Sym}(L)$ can be alternatively defined in terms of L and K alone with no reference to any linear mapping A as the next proposition states.

Proposition 6. Assume $L \in \text{Gr}(V, m)$. Then

$$\text{Sym}(L) := \min_{\substack{v \in K \\ \|v\|=1}} \max_{x \in K, t \in \mathbb{R}} \{t : x + tv \in L, \|x\| \leq 1\}.$$

Proof. Assume $A : V \rightarrow W$ is such that $L = \ker(A)$. From (13) and (14) it follows that for $S := \{Ax : x \in K, \|x\| \leq 1\}$

$$\begin{aligned} \text{Sym}(L) &= \min_{\substack{v \in K \\ \|v\| \leq 1}} \max_{t \in \mathbb{R}} \{t : -tAv \in S\} \\ &= \min_{\substack{v \in K \\ \|v\| \leq 1}} \max_{x \in K, t \in \mathbb{R}} \{t : -tAv = Ax, \|x\| \leq 1\} \\ &= \min_{\substack{v \in K \\ \|v\| \leq 1}} \max_{x \in K, t \in \mathbb{R}} \{t : x + tv \in L, \|x\| \leq 1\}. \end{aligned}$$

□

Observe that $L \cap \text{int}(K) \neq \emptyset$ if and only if $\text{Sym}(L) > 0$. It is also easy to see that $\text{Sym}(L) \in [0, 1]$ for all $L \in \text{Gr}(V, m)$. The following result is a general version of [15, Proposition 22].

Theorem 3. *Assume $L \in \text{Gr}(V, m)$ is such that $L \cap \text{int } K \neq \emptyset$, $\|\cdot\| = \|\cdot\|_e$ and $\text{Sym}(L) < 1$. Then*

$$\frac{\text{Sym}(L)}{1 + \text{Sym}(L)} \leq \sigma(L) \leq \frac{\text{Sym}(L)}{1 - \text{Sym}(L)}.$$

If $\|\cdot\| = \|\cdot\|_e^*$ for some $e \in \text{int}(K^*)$ then

$$\frac{\text{Sym}(L)}{1 + \text{Sym}(L)} = \sigma(L).$$

Proof. To ease notation, let $s := \text{Sym}(L)$ and $\sigma := \sigma(L)$. First we show that $\sigma \geq \frac{s}{1+s}$. To that end, assume $v \in K, \|v\| = 1$ is given. It follows from the definition of $\text{Sym}(L)$ that there exists $z \in K, \|z\| \leq 1$ such that $z + sv \in L$. Thus $x := \frac{1}{\|z+sv\|}(z + sv) \in L, \|x\| = 1$ and

$$\lambda_v(x) \geq \frac{s}{\|z + sv\|} \geq \frac{s}{1 + s}.$$

Since this holds for any $v \in K, \|v\| = 1$, it follows that $\sigma \geq \frac{s}{1+s}$.

Next we show that $\sigma \leq \frac{s}{1-s}$. To that end, assume $v \in K, \|v\| = 1$ is such that

$$\max_{x \in K, t \in \mathbb{R}} \{t : x + tv \in L, \|x\| \leq 1\} < 1. \quad (15)$$

At least one such v exists because $\text{Sym}(L) < 1$.

It follows from the construction of $\sigma(L)$ that there exists $x \in L, \|x\| = 1$ such that $\lambda_v(x) = \sigma > 0$. In particular, $x - \sigma v \in K$. Furthermore, $x - \sigma v \neq 0$ as otherwise $v = \frac{1}{\sigma}x \in L$ and $x + v \in L$ which would contradict (15). Thus $z := \frac{x - \sigma v}{\|x - \sigma v\|} \in K, \|z\| = 1$ and $z + \frac{\sigma}{\|x - \sigma v\|}v \in L$ with $\frac{\sigma}{\|x - \sigma v\|} \geq \frac{\sigma}{1 + \sigma}$. Since this holds for any $v \in K, \|v\| = 1$ satisfying (15), it follows that $s \geq \frac{\sigma}{1 + \sigma}$ or equivalently $\sigma \leq \frac{s}{1 - s}$.

In the special case when $\|\cdot\| = \|\cdot\|_e^*$ for some $e \in \text{int}(K^*)$ we have $\|z\| = \langle e, z \rangle$ for all $z \in K$. In particular, $\|x - \sigma v\| = \langle e, x - \sigma v \rangle = \langle e, x \rangle - \langle e, \sigma v \rangle = \|x\| - \sigma\|v\| = 1 - \sigma$ in the previous paragraph and so the second inequality can be sharpened to $s \geq \frac{\sigma}{1 - \sigma}$ or equivalently $\sigma \leq \frac{s}{1 + s}$. \square

We also have the following relationship between the Grassmann distance to ill-posedness and the symmetry measure.

Corollary 2. Assume $L \in \text{Gr}(V, m)$ is such that $L \cap \text{int}(K) \neq \emptyset$ and $\|\cdot\| = \|\cdot\|$. Then

$$\min_{\substack{u \in K^* \\ \|u\|^*=1}} \max_{\substack{v \in K \\ \|v\|=1}} \langle u, v \rangle \cdot \frac{\text{Sym}(L)}{1 + \text{Sym}(L)} \leq \text{dist}_G(L, \Sigma_m) \leq \frac{\text{Sym}(L)}{1 - \text{Sym}(L)}.$$

In particular, if $\|\cdot\| = \|\cdot\| = \|\cdot\|_2$ then

$$\cos(\Theta(K^*, K)) \cdot \frac{\text{Sym}(L)}{1 + \text{Sym}(L)} \leq \text{dist}_G(L, \Sigma_m) \leq \frac{\text{Sym}(L)}{1 - \text{Sym}(L)}.$$

Proof. This is an immediate consequence of Theorem 1, Theorem 3, and Corollary 1. \square

5 Symmetric and polyhedral cones

In this section we specialize our developments to two important cases. The first one is when K is a symmetric cone and the space V is endowed with an Euclidean Jordan algebra structure. The second one is the case $V = \mathbb{R}^n$, $K = \mathbb{R}_+^n$.

5.1 Symmetric cones

We next recall some basic material on Euclidean Jordan algebras. The articles [27, 28] and the textbooks [4, 17] provide a more detailed discussion on this topic and its relevance to optimization.

Assume V is endowed with a bilinear operation $\circ : V \times V \rightarrow V$ and $e \in V$ is a particular element of V . The triple (V, \circ, e) is an *Euclidean Jordan algebra with identity element* if the following conditions hold:

- $x \circ y = y \circ x$ for all $x, y \in V$
- $x \circ (x^2 \circ y) = x^2 \circ (x \circ y)$ for all $x, y \in V$, where $x^2 = x \circ x$
- $x \circ e = x$ for all $x \in V$
- There exists an associative positive definite bilinear form on V .

An element $c \in V$ is *idempotent* if $c^2 = c$. An idempotent element of V is a *primitive idempotent* if it is not the sum of two other idempotents. We will write $c \in \text{PI}$ as shorthand for the statement ‘ c is a primitive idempotent.’ The rank r of V is the smallest integer such that for all $x \in V$ the set

$\{e, x, x^2, \dots, x^r\}$ is linearly dependent. Every element $x \in V$ has a *spectral decomposition*

$$x = \sum_{i=1}^r \lambda_i(x) c_i, \quad (16)$$

where $\lambda_i(x) \in \mathbb{R}$, $i = 1, \dots, r$ are the *eigenvalues* of x and $\{c_1, \dots, c_r\}$ is a *Jordan frame*, that is, a collection of non-zero primitive idempotents such that $c_i \circ c_j = 0$ for $i \neq j$, and $c_1 + \dots + c_r = e$. The *trace* of $x \in V$ is defined as

$$\text{trace}(x) = \sum_{i=1}^r \lambda_i(x)$$

Throughout this section we assume that (V, \circ, e) is an Euclidean Jordan algebra with identity and let r denote the rank of V . Furthermore, we assume that V is endowed with the following trace inner product:

$$\langle x, y \rangle := \text{trace}(x \circ y). \quad (17)$$

We also assume that K is the cone of squares in V , that is, $K = \{x^2 : x \in V\}$. It is known that K is a *symmetric cone*, that is, $K = K^*$ and for all $u, v \in \text{int}(K)$ there exists a linear automorphism $A \in GL(V)$ such that $Au = v$ and $AK = K$.

The following four cases summarize the main examples of Euclidean Jordan algebras that are popular in optimization. First, \mathbb{R}^n with the componentwise product $(x \circ z)_i = x_i z_i$, $i = 1, \dots, n$. In this case $K = \mathbb{R}_+^n$ and $r = n$. Second, \mathbb{S}^n with the product $X \circ Z = \frac{XZ + ZX}{2}$. In this case $K = \mathbb{S}_+^n$ and $r = n$. Third, \mathbb{R}^n with the product $\begin{bmatrix} x_0 \\ \bar{x} \end{bmatrix} \circ \begin{bmatrix} z_0 \\ \bar{z} \end{bmatrix} = \begin{bmatrix} x^\top z \\ x_0 \bar{z} + z_0 \bar{x} \end{bmatrix}$. In this case $K = \{(x_0, \bar{x}) \in \mathbb{R}^n : x_0 \geq \|\bar{x}\|_2\}$ and $r = 2$. Fourth, any direct product of the above types of Euclidean Jordan algebras is also an Euclidean Jordan algebra.

For $x \in V$, the vector of eigenvalues $\lambda(x) := (\lambda_1(x), \dots, \lambda_r(x))$ is nicely related to the norms $\|x\|_2, \|x\|_e, \|x\|_e^*$ as follows

$$\|x\|_2 = \|\lambda(x)\|_2, \quad \|x\|_e = \|\lambda(x)\|_\infty, \quad \|x\|_e^* = \|\lambda(x)\|_1. \quad (18)$$

It is also easy to see that the eigenvalue mapping λ_e satisfies $\lambda_e(x) = \min_{i=1, \dots, r} \lambda_i(x)$ and the set $\Delta(K)$ can be written in any of the following ways

$$\{x \in K : \|x\|_e^* = 1\} = \{x \in K : \langle e, x \rangle = 1\} = \left\{ x \in K : \sum_{i=1}^r \lambda_i(x) = 1 \right\}.$$

Part (a) in Proposition 7 below can be seen as an extension of the identities (9) and (10) in Section 3.

Proposition 7. *Assume $L \in \text{Gr}(V, m)$ is such that $L \cap \text{int}(K) \neq \emptyset$.*

(a) *If $\| \cdot \| = \| \cdot \|_e$ then*

$$\sigma(L) = \nu(L) = \max_{\substack{x \in L \cap K \\ \|x\|=1}} \min_{i \in 1, \dots, r} \lambda_i(x).$$

In particular, if $\| \cdot \| = \| \cdot \|_e^$ then*

$$\sigma(L) = \nu(L) = \max_{x \in L \cap \Delta(K)} \min_{i \in 1, \dots, r} \lambda_i(x).$$

(b) *If $\| \cdot \| = \| \cdot \|_e^*$ then*

$$\sigma(L) = \nu(L) = \min_{c \in \text{PI}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \langle c, x \rangle.$$

In particular, if $\| \cdot \| = \| \cdot \|_e^$ then*

$$\sigma(L) = \nu(L) = \min_{c \in \text{PI}} \max_{x \in L \cap \Delta(K)} \langle c, x \rangle.$$

(Recall that $c \in \text{PI}$ is shorthand for the statement ‘ c is a primitive idempotent.’)

(c) *If $\| \cdot \| = \| \cdot \|_2$ then*

$$\sigma(L) = \nu(L) = \min_{\substack{u \in K \\ \|u\|_2=1}} \max_{\substack{x \in L \cap K \\ \|x\|_2=1}} \langle u, x \rangle.$$

Proof. (a) Since $L \cap \text{int}(K) \neq \emptyset$ and $\| \cdot \| = \| \cdot \|_e$, Proposition 4 and Corollary 1(c) yield

$$\sigma(L) = \nu(L) = \max_{\substack{x \in L \\ \|x\| \leq 1}} \lambda_e(x) = \max_{\substack{x \in L \cap K \\ \|x\|=1}} \min_{i=1, \dots, r} \lambda_i(x).$$

If in addition $\| \cdot \| = \| \cdot \|_e^*$ then $\|x\| = \langle e, x \rangle$ for all $x \in K$. In other words, $\{x \in K : \|x\| = 1\} = \Delta(K)$. Thus

$$\sigma(L) = \nu(L) = \max_{\substack{x \in L \cap K \\ \|x\|=1}} \min_{i=1, \dots, r} \lambda_i(x) = \max_{x \in L \cap \Delta(K)} \min_{i=1, \dots, r} \lambda_i(x).$$

- (b) The first statement is a consequence of the following inequalities, which by putting them together must hold with equality:

$$\sigma(L) \leq \min_{c \in \text{PI}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \lambda_c(x) \leq \min_{c \in \text{PI}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \langle c, x \rangle, \quad (19)$$

$$\nu(L) \geq \min_{\substack{u \in K \\ \|u\|_e=1}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \langle u, x \rangle = \min_{c \in \text{PI}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \langle c, x \rangle, \quad (20)$$

$$\sigma(L) \geq \nu(L). \quad (21)$$

To prove (19), observe that if $c \in \text{PI}$ then $\langle c, c \rangle = 1$ and consequently $\lambda_c(x) \leq \langle c, x \rangle$ for all $x \in L$. Furthermore, if $c \in \text{PI}$ then $\|c\|_e^* = \langle e, c \rangle = 1$. Therefore from (11) we get

$$\sigma(L) = \min_{\substack{v \in K \\ \langle e, v \rangle=1}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \lambda_v(x) \leq \min_{c \in \text{PI}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \lambda_c(x) \leq \min_{c \in \text{PI}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \langle c, x \rangle.$$

To prove (20), observe that if $u \in K$ and $\|u\|_e = 1$ then, from the spectral decomposition of u and (18), it follows that $u - c \in K$ for some $c \in \text{PI}$ and so $\langle u, x \rangle \geq \langle c, x \rangle$ for all $x \in L \cap K$. Thus Proposition 2 implies

$$\nu(L) = \min_{\substack{u \in K \\ \|u\|_e=1}} \max_{\substack{x \in L \\ \|x\|=1}} \langle u, x \rangle \geq \min_{\substack{u \in K \\ \|u\|_e=1}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \langle u, x \rangle = \min_{c \in \text{PI}} \max_{\substack{x \in L \cap K \\ \|x\|=1}} \langle c, x \rangle.$$

Finally (21) follows from Corollary 1(a) and Hölder's inequality (4).

The second statement follows as in part (a).

- (c) The reasoning is almost identical to that in part (b). This is a consequence of the following inequalities whose proofs are straightforward modifications of the proofs of (19), (20), and (21):

$$\sigma(L) = \min_{\substack{u \in K \\ \|u\|_2=1}} \max_{\substack{x \in L \cap K \\ \|x\|_2=1}} \lambda_u(x) \leq \min_{\substack{u \in K \\ \|u\|_2=1}} \max_{\substack{x \in L \cap K \\ \|x\|_2=1}} \langle u, x \rangle, \quad (22)$$

$$\nu(L) = \min_{\substack{u \in K \\ \|u\|_2=1}} \max_{\substack{x \in L \\ \|x\|_2=1}} \langle u, x \rangle \geq \min_{\substack{u \in K \\ \|u\|_2=1}} \max_{\substack{x \in L \cap K \\ \|x\|_2=1}} \langle u, x \rangle, \quad (23)$$

$$\sigma(L) \geq \nu(L). \quad (24)$$

□

Corollary 3 below is an immediate consequence of Proposition 7. Corollary 3(b) shows that if the problem

$$\text{find } w \text{ such that } Aw \in \text{int}(K)$$

is feasible, then it can be recast as the *preconditioned* equivalent problem

$$\text{find } w \text{ such that } (PAR)w \in \text{int}(K)$$

for suitable linear automorphisms P, R where the latter reformulation is well-conditioned.

Corollary 3. Assume $\|\cdot\| = \|\cdot\|_2$.

- (a) If $L \in \text{Gr}(V, m)$ is such that $L \cap \text{int}(K) \neq \emptyset$ then there exists a linear automorphism $P \in GL(V)$ such that $PK = K$ and

$$\nu(PL) \geq \frac{1}{\sqrt{r}}.$$

- (b) Assume $A : W \rightarrow V$ is injective and $\text{Im}(A) \cap \text{int}(K) \neq \emptyset$. Then there exist linear automorphisms $P \in GL(V), R \in GL(W)$ such that $PK = K$ and

$$\frac{\|PAR\|}{\text{dist}_R(PAR, \Sigma_m)} \leq \sqrt{r}.$$

Proof. (a) Let $x_0 \in L \cap \text{int}(K)$. Since K is symmetric, there exists $P \in GL(V)$ such that $PK = K$ and $Px_0 = e$. Observe that $e \in PL \cap \text{int}(K)$, $\|e\|_2 = \sqrt{r}$, and $\langle e, u \rangle = \|u\|_e \geq \|u\|_2$ for all $u \in K$. Thus Proposition 7(c) implies that $\nu(PL) \geq \frac{1}{\sqrt{r}}$.

- (b) Let $L = \text{Im}(A)$ and P be as in part (a). Let $R \in GL(W)$ be such that PAR is balanced, that is, $(PAR)^*(PAR) = I$. Thus, by Theorem 2,

$$\frac{\|PAR\|}{\text{dist}_R(PAR, \Sigma_m)} = \frac{1}{\nu(PL)} \leq \sqrt{r}.$$

□

5.2 Polyhedral case

We now consider the special case $V = \mathbb{R}^n, K = \mathbb{R}_+^n$. The peculiar structure of \mathbb{R}_+^n enables a refinement of the the former condition measures to all $L \in \text{Gr}(\mathbb{R}^n, m)$ including Σ_m via a suitable canonical partition of $\{1, \dots, n\}$.

More precisely, the classical Goldman-Tucker partition theorem implies that for all $L \in \text{Gr}(\mathbb{R}^n, m)$ there exists a unique partition $B \cup N = \{1, \dots, n\}$ such that

$$L \cap (\text{int } \mathbb{R}_+^B \times \{0_N\}) \neq \emptyset \text{ and } L^\perp \cap (\{0_B\} \times \text{int } \mathbb{R}_+^N) \neq \emptyset. \quad (25)$$

(A detailed proof of the existence of this partition can be found in [11].)

Note that $L \notin \Sigma$ if and only if either $B = \emptyset$ or $N = \emptyset$. Regardless of whether $L \in \Sigma_m$, the partition B, N yields a decomposition into two subproblems each of which is well-posed as we detail next.

Let $V_B := \mathbb{R}^B \times \{0\}$, $K_B := \mathbb{R}_+^B \times \{0\}$, $V_N := \{0\} \times \mathbb{R}^N$, $K_N := \{0\} \times \mathbb{R}_+^N$ and $L_B := L \cap V_B$, $L_N = L^\perp \cap V_N$. Observe that relative to the space-cone pairs (V_B, K_B) and (V_N, K_N) both L_B and L_N are respectively well-posed since $L_B \cap \text{int}(K_B) \neq \emptyset$ and $L_N \cap \text{int}(K_N) \neq \emptyset$. Therefore, the partition (B, N) yields a natural decomposition of the entire previous machinery. More precisely, we have

$$\nu(L_B) = \min_{\substack{u \in K_B, y \in L_B^\perp \\ \|u\| = 1}} \|u - y\|^*, \quad \nu(L_N) = \min_{\substack{u \in K_N, y \in L_N^\perp \\ \|u\| = 1}} \|u - y\|^*, \quad (26)$$

and

$$\sigma(L_B) = \min_{\substack{v \in K_B \\ \|v\| = 1}} \max_{\substack{x \in L_B \\ \|x\| \leq 1}} \lambda_v(x), \quad \sigma(L_N) = \min_{\substack{v \in K_N \\ \|v\| = 1}} \max_{\substack{x \in L_N \\ \|x\| \leq 1}} \lambda_v(x). \quad (27)$$

Furthermore, if $k = \dim(L_B)$ and $\ell = \dim(L_N)$ then

$$\text{dist}_G(L_B, \Sigma_{k,B}) = \nu(L_B), \quad \text{dist}_G(L_N, \Sigma_{\ell,N}) = \nu(L_N)$$

where $\Sigma_{k,B} \subseteq \text{Gr}(V_B, k)$ and $\Sigma_{\ell,N} \subseteq \text{Gr}(V_N, \ell)$ are respectively the set of k -dimensional and ℓ -dimensional ill-posed subspaces in V_B and V_N respectively.

The following result readily follows from Proposition 7 specialized to each of the two subproblems defined by the partition (B, N) . Proposition 8(b) below shows that $\sigma(L_B)$ and $\sigma(L_N)$ coincide with the sigma measure introduced by Ye [29].

Proposition 8. *Assume $L \in \text{Gr}(\mathbb{R}^n, m)$ and let (B, N) is the partition determined by L as in (25).*

(a) *If $\| \cdot \| = \| \cdot \|_\infty$ then*

$$\sigma(L_B) = \nu(L_B) = \max_{\substack{x \in L \cap \mathbb{R}_+^n \\ \|x\| = 1}} \min_{i \in B} x_i, \quad \sigma(L_N) = \nu(L_N) = \max_{\substack{x \in L^\perp \cap \mathbb{R}_+^n \\ \|x\| = 1}} \min_{i \in N} x_i$$

In particular, if $\|\cdot\| = \|\cdot\|_1$ then

$$\sigma(L_B) = \nu(L_B) = \max_{x \in L \cap \Delta_{n-1}} \min_{i \in B} x_i, \quad \sigma(L_N) = \nu(L_N) = \max_{x \in L^\perp \cap \Delta_{n-1}} \min_{i \in N} x_i.$$

(b) If $\|\cdot\| = \|\cdot\|_1$ then

$$\sigma(L_B) = \nu(L_B) = \min_{i \in B} \max_{\substack{x \in L \cap \mathbb{R}_+^n \\ \|x\|=1}} x_i, \quad \sigma(L_N) = \nu(L_N) = \min_{i \in N} \max_{\substack{x \in L^\perp \cap \mathbb{R}_+^n \\ \|x\|=1}} x_i$$

In particular, if $\|\cdot\| = \|\cdot\|_1$ then

$$\sigma(L_B) = \nu(L_B) = \min_{i \in B} \max_{x \in L \cap \Delta_{n-1}} x_i, \quad \sigma(L_N) = \nu(L_N) = \min_{i \in N} \max_{x \in L^\perp \cap \Delta_{n-1}} x_i.$$

(c) If $\|\cdot\| = \|\cdot\|_2$ then

$$\sigma(L_B) = \nu(L_B) = \min_{\substack{u \in \mathbb{R}_+^B \times \{0_N\} \\ \|u\|_2=1}} \max_{\substack{x \in L \cap \mathbb{R}_+^n \\ \|x\|_2=1}} \langle u, x \rangle,$$

and

$$\sigma(L_N) = \nu(L_N) = \min_{\substack{u \in \{0_B\} \times \mathbb{R}_+^N \\ \|u\|_2=1}} \max_{\substack{x \in L^\perp \cap \mathbb{R}_+^n \\ \|x\|_2=1}} \langle u, x \rangle.$$

We conclude with a discussion analogous to that in Section 2.2. Assume $A : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is an injective linear mapping and $L = \text{Im}(A)$. The partition (B, N) determined by L in turn yields the following decomposition of A and \mathbb{R}^m . First, the mapping A decomposes as $A = \begin{bmatrix} A_B \\ A_N \end{bmatrix}$. Second, the space \mathbb{R}^m has the orthogonal decomposition $\mathbb{R}^m = \mathbb{R}_B^m + \mathbb{R}_N^m$ where $\mathbb{R}_B^m := \text{Im}(A_B^\top)$ and $\mathbb{R}_N^m := \ker(A_B) = (\mathbb{R}_B^m)^\perp$. Consequently, the mapping A can be written in the following block matrix form

$$A = \begin{bmatrix} A_B \\ A_N \end{bmatrix} = \begin{bmatrix} A_{BB} & 0 \\ A_{NB} & A_{NN} \end{bmatrix},$$

where A_{BB} and A_{NN} are both full column-rank and the spaces L_B and L_N are

$$L_B = \text{Im}(A_{BB}) \times \{0_N\}, \quad L_N = \{0_B\} \times \ker(A_{NN}^\top).$$

Theorem 2 readily yields

$$\frac{1}{\|A_{BB}\|} \leq \frac{\nu(L_B)}{\text{dist}_R(A_{BB}, \Sigma)} \leq \|A_{BB}^{-1}\|,$$

and

$$\frac{1}{\|A_{NN}\|} \leq \frac{\bar{\nu}(L_N^\perp)}{\text{dist}_R(A_{NN}, \Sigma)} \leq \|A_{NN}^{-1}\|.$$

Furthermore, proceeding as in Corollary 3, it follows that if $\|\cdot\| = \|\cdot\| = \|\cdot\|_2$, then for a suitable positive diagonal matrix D and non-singular matrix R the space $DL = \text{Im}(\hat{A})$ of the preconditioned matrix $\hat{A} = DAR$ determines the same partition (B, N) as $L = \text{Im}(A)$ and the preconditioned matrix \hat{A} satisfies

$$\frac{\|\hat{A}_{BB}\|}{\text{dist}_R(\hat{A}_{BB}, \Sigma)} = \frac{1}{\nu((DL)_B)} \leq \sqrt{k},$$

and

$$\frac{\|\hat{A}_{NN}\|}{\text{dist}_R(\hat{A}_{NN}, \Sigma)} = \frac{1}{\nu((DL)_N)} \leq \sqrt{\ell}.$$

References

- [1] D. Amelunxen and P. Bürgisser. A coordinate-free condition number for convex programming. *SIAM J. Optim.*, 22(3):1029–1041, 2012.
- [2] D. Amelunxen and P. Bürgisser. Probabilistic analysis of the Grassmann condition number. *Foundations of Computational Mathematics*, 15(1):3–51, 2015.
- [3] D. Amelunxen, M. Lotz, M. McCoy, and J. Tropp. Living on the edge: phase transitions in convex programs with random data. *Inf. Inference*, 3(3):224–294, 2014.
- [4] M. Baes. *Spectral Functions and Smoothing Techniques on Jordan Algebras: How algebraic techniques can help to design efficient optimization algorithms*. Lambert Academic Publishing, 2009.
- [5] A. Belloni and R. Freund. On the symmetry function of a convex set. *Math. Program.*, 111(1-2, Ser. B):57–93, 2008.
- [6] A. Belloni and R. Freund. Projective re-normalization for improving the behavior of a homogeneous conic linear system. *Math. Program.*, 118(2, Ser. A):279–299, 2009.
- [7] A. Belloni and R. M. Freund. A geometric analysis of Renegar’s condition number, and its interplay with conic curvature. *Math. Program.*, 119(1, Ser. A):95–107, 2009.

- [8] D. Bertsimas, V. Goyal, and X. Sun. A geometric characterization of the power of finite adaptability in multistage stochastic and adaptive optimization. *Math. Oper. Res.*, 36(1):24–54, 2011.
- [9] J. Borwein and A. Lewis. *Convex analysis and nonlinear optimization*. CMS Books in Mathematics/Ouvrages de Mathématiques de la SMC, 3. Springer, New York, second edition, 2006. Theory and examples.
- [10] D. Cheung and F. Cucker. A new condition number for linear programming. *Math. Prog.*, 91(2):163–174, 2001.
- [11] D. Cheung, F. Cucker, and J. Peña. Unifying condition numbers for linear programming. *Math. Oper. Res.*, 28(4):609–624, 2003.
- [12] D. Cheung, F. Cucker, and J. Peña. A condition number for multifold conic systems. *SIAM J. Optim.*, 19(1):261–280, 2008.
- [13] D. Cheung, F. Cucker, and J. Peña. On strata of degenerate polyhedral cones. II. Relations between condition measures. *J. Complexity*, 26(2):209–226, 2010.
- [14] A. L. Dontchev, A. S. Lewis, and R. T. Rockafellar. The radius of metric regularity. *Trans. Amer. Math. Soc.*, 355(2):493–517 (electronic), 2003.
- [15] M. Epelman and R. Freund. A new condition measure, preconditioners, and relations between different measures of conditioning for conic linear systems. *SIAM J. Optim.*, 12(3):627–655 (electronic), 2002.
- [16] M. Epelman and R. M. Freund. Condition number complexity of an elementary algorithm for computing a reliable solution of a conic linear system. *Math Program.*, 88(3):451–485, 2000.
- [17] J. Faraut and A. Korányi. *Analysis on Symmetric Cones*. Oxford Mathematical Monographs, 1994.
- [18] R. Freund. Complexity of convex optimization using geometry-based measures and a reference point. *Math Program.*, 99:197–221, 2004.
- [19] R. Freund and J. Vera. Condition-based complexity of convex optimization in conic linear form via the ellipsoid algorithm. *SIAM J. on Optim.*, 10:155–176, 1999.
- [20] R. Freund and J. Vera. Some characterizations and properties of the “distance to ill-posedness” and the condition measure of a conic linear system. *Math Program.*, 86:225–260, 1999.

- [21] G. Golub and C. Van Loan. *Matrix Computations*. Johns Hopkins University Press, Baltimore, third edition, 1996.
- [22] A. Lewis. Ill-conditioned convex processes and conic linear systems. *Math. Oper. Res.*, 24(4):829–834, 1999.
- [23] J. Peña. Understanding the geometry on infeasible perturbations of a conic linear system. *SIAM J. on Optim.*, 10:534–550, 2000.
- [24] J. Peña. Conic systems and sublinear mappings: equivalent approaches. *Oper. Res. Lett.*, 32(5):463–467, 2004.
- [25] J. Renegar. Incorporating condition measures into the complexity theory of linear programming. *SIAM J. on Optim.*, 5:506–524, 1995.
- [26] J. Renegar. Linear programming, complexity theory and elementary functional analysis. *Math. Programming*, 70(3, Ser. A):279–351, 1995.
- [27] S. Schmieta and F. Alizadeh. Associative and Jordan algebras, and polynomial time interior-point algorithms for symmetric cones. *Math. Oper. Res.*, 26(3):543–564, 2001.
- [28] S. Schmieta and F. Alizadeh. Extension of primal-dual interior point algorithms to symmetric cones. *Math Program.*, 96(3):409–438, 2003.
- [29] Y. Ye. Toward probabilistic analysis of interior-point algorithms for linear programming. *Math. of Oper. Res.*, 19:38–52, 1994.